Plotting with ggplot2

Part II
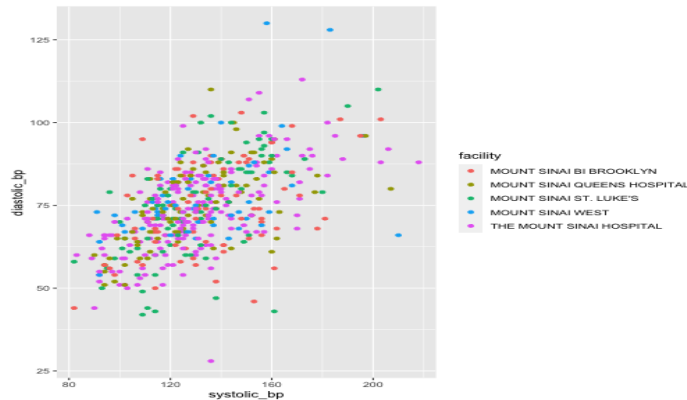
Joselyn Chávez

05/24/2022
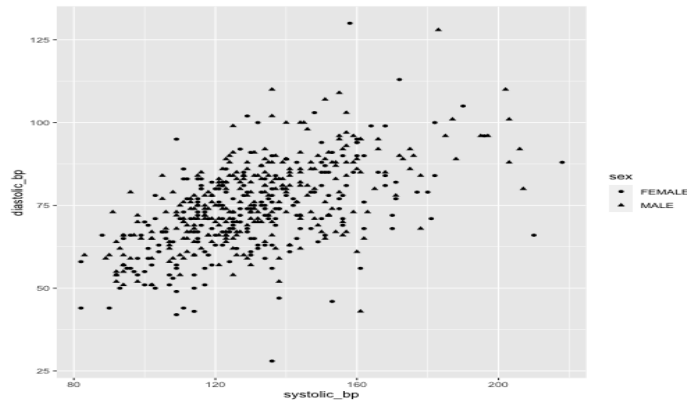
Plotting 3 variables (num-num-cat)

```
ggplot(sinai_covid,
       aes(x = systolic_bp,
           y = diastolic_bp,
           color = facility)) +
  geom_point()
```

Plotting 3 variables (num-num-cat)

```
ggplot(sinai_covid,
       aes(x = systolic_bp,
           y = diastolic_bp,
           shape = sex)) +
  geom_point()
```
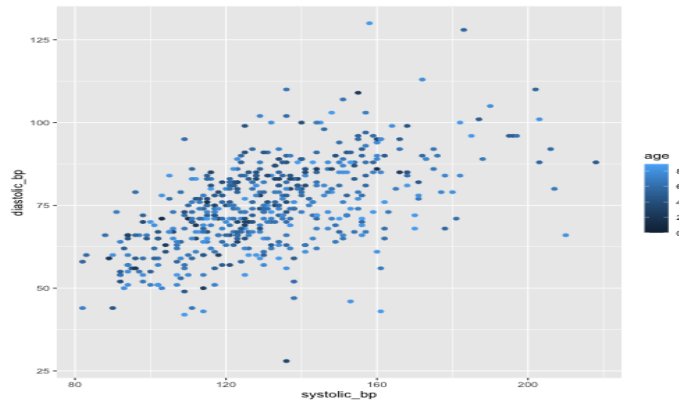
Let's practice

- Plot the patient's age vs bmi, and separate by color and/or shape based on their smoking status

Plotting 3 variables (num-num-num)

```
ggplot(sinai_covid,
       aes(x = systolic_bp,
           y = diastolic_bp,
           color = age)) +
  geom_point()
```
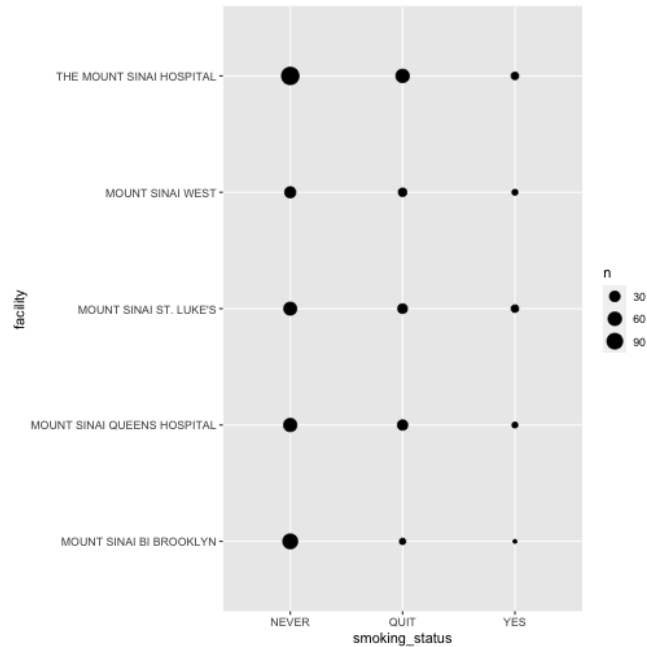
Plotting 3 variables (cat-cat-num)

```
# first create the summarized tibble
count_smoking <- sinai_covid %>%
  group_by(facility) %>%
  count(smoking_status)

# plot
ggplot(count_smoking,
       aes(x = smoking_status,
           y = facility,
           size = n)) +
  geom_point()
```
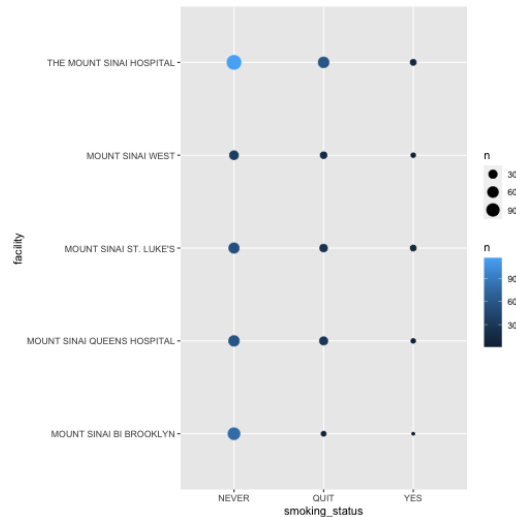
Plotting 3 variables (cat-cat-num)

Plotting 3 variables (cat-cat-num)

```
ggplot(count_smoking,
       aes(x = smoking_status,
           y = facility,
           size = n, color = n)) +
  geom_point()
```
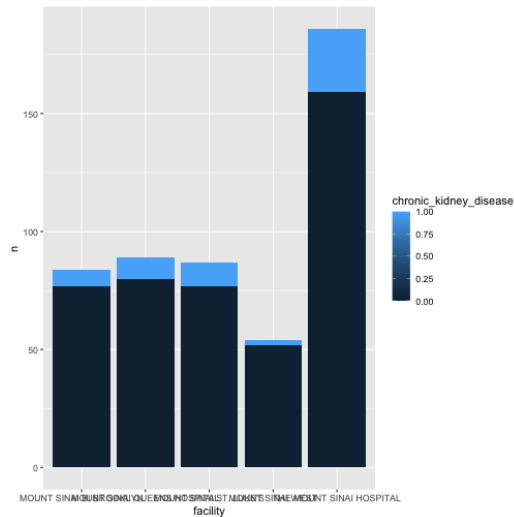
Let's practice

- Count the number of patients with chronic kidney disease (0 and 1 separately) per facility
- Plot the data using a dot plot, mapping the size of the dots to the chronic kidney disease status
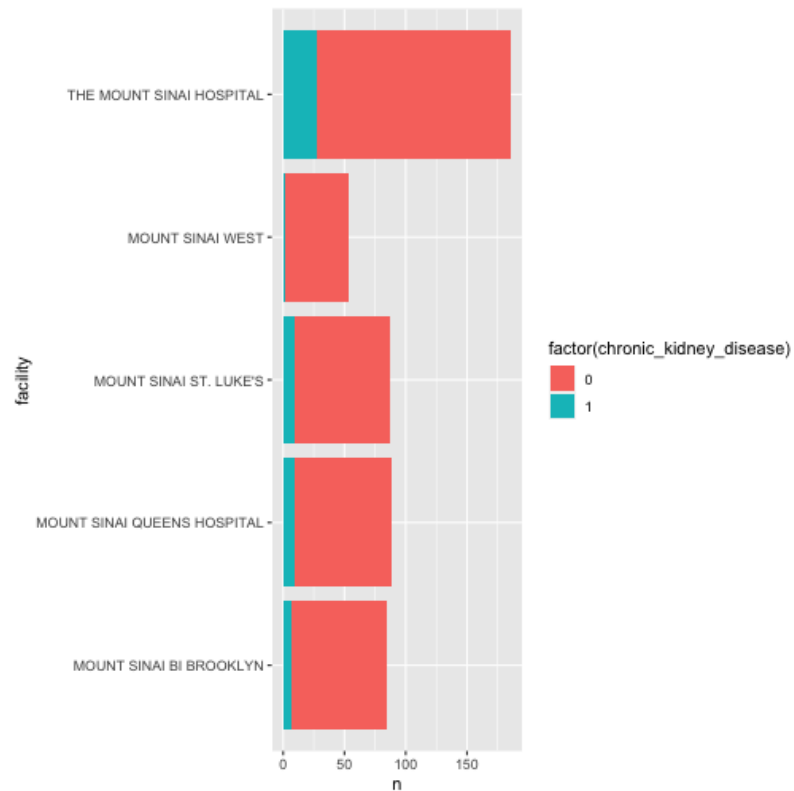
Another way

```
kd <- sinai_covid %>%
  group_by(facility) %>%
  count(chronic_kidney_disease)

ggplot(kd, aes(x = facility, y = n,
               fill = chronic_kidney_disease)) +
  geom_col()
```
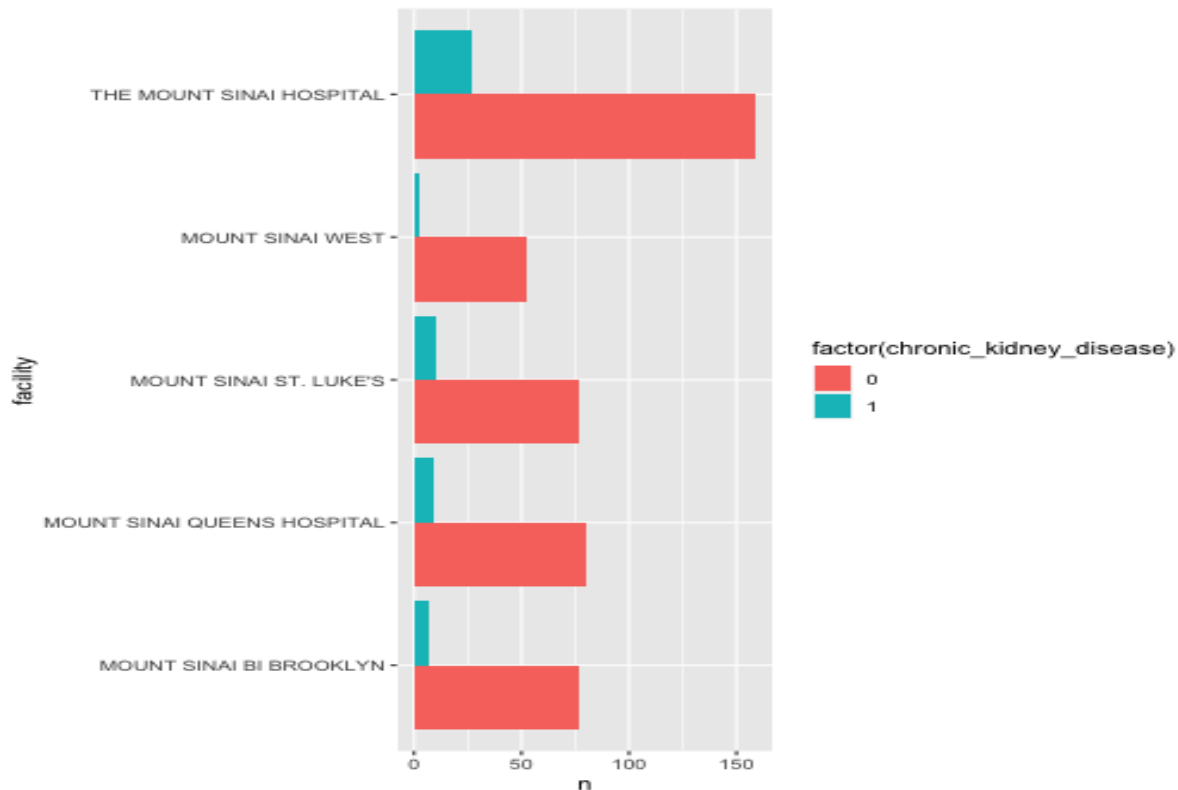
Let's treat the variable as factor

```
ggplot(kd, aes(x = n, y = facility,
               fill = factor(chronic_kidney_disease))) +
  geom_col()
```
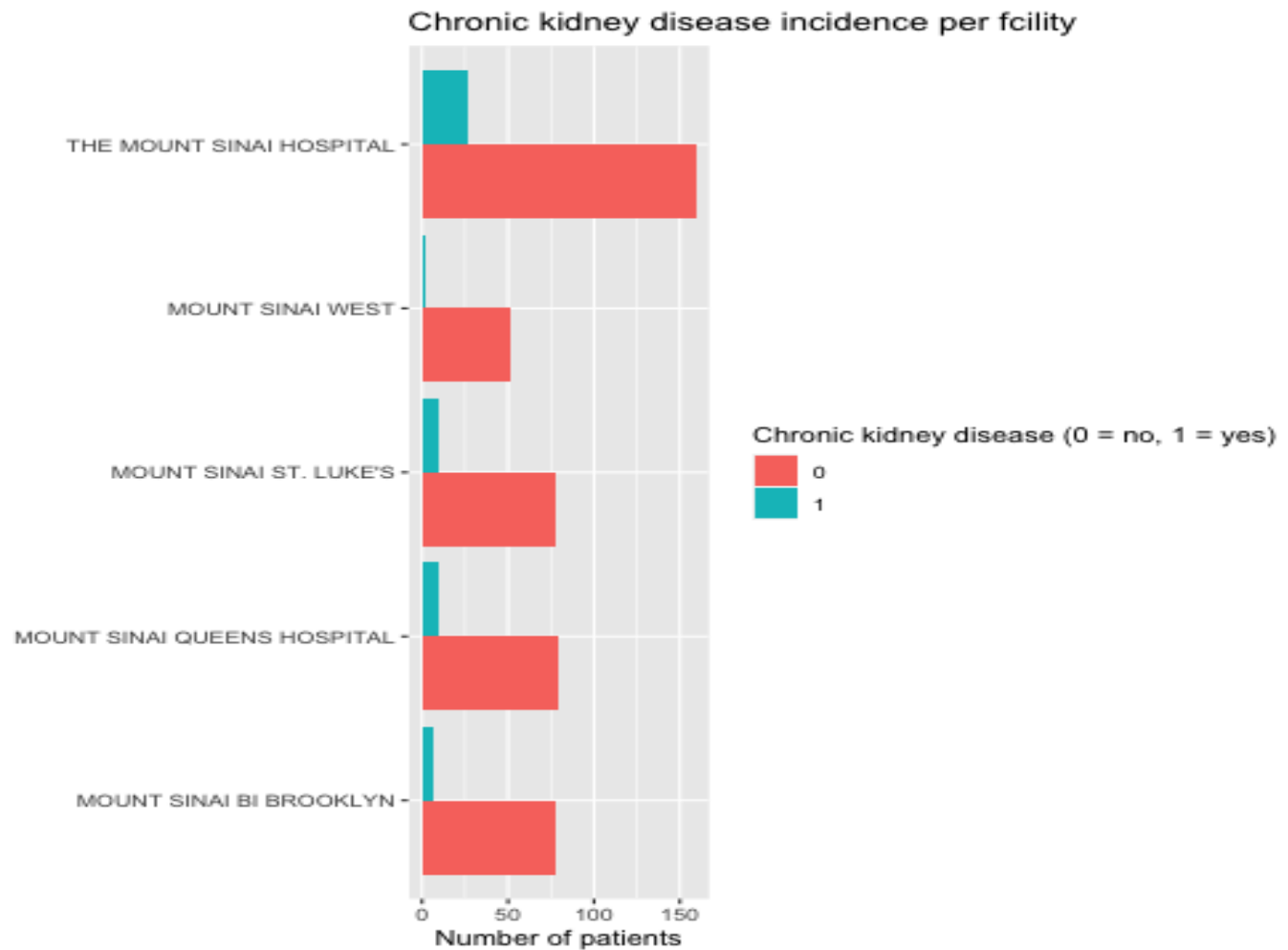
Splitting the bars

```
ggplot(kd, aes(x = n, y = facility,
          fill = factor(chronic_kidney_disease))) +
  geom_col(position = position_dodge())
```

Adding titles

```
ggplot(kd,
       aes(x = n, y = facility,
           fill = factor(chronic_kidney_disease))) +
  geom_col(position = position_dodge()) +
  labs(title = "Chronic kidney disease incidence per fcility",
       x = "Number of patients",
       y = "",
       fill = "Chronic kidney disease (0 = no, 1 = yes)")
```

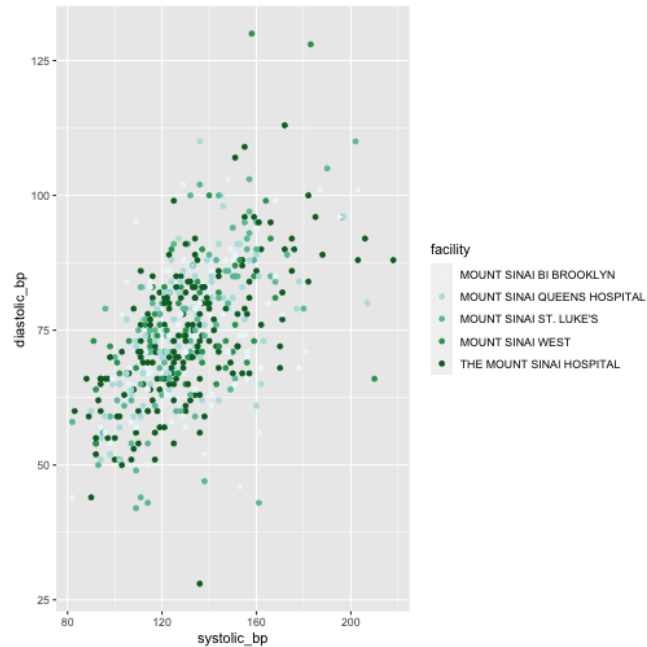Chronic kidney disease incidence per fcility

Let's practice!

- Count the number of patients depending on their smoking status per ethnic group
- Represent the data on a column plot
- Split each bar depending on the smoking status
- Add a plot title, axis title and change the legend title to remove the underscores
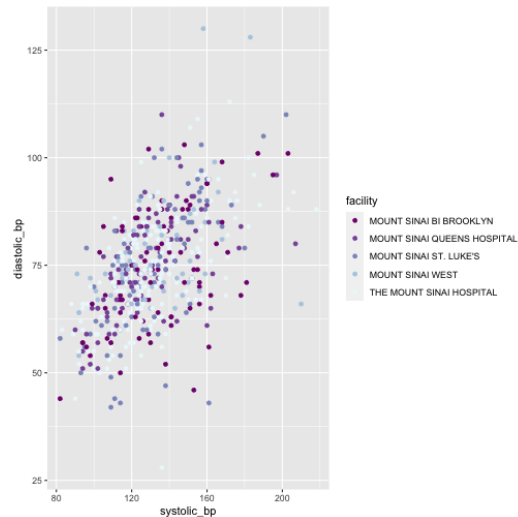
Scales

```
ggplot(sinai_covid, aes(x = systolic_bp, y = diastolic_bp,
          color = facility)) +
  geom_point() +
  scale_color_brewer(palette = 2)
```
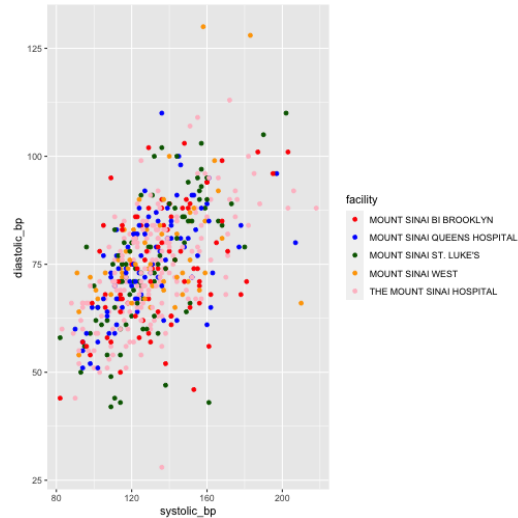
Scales

- Invert scale direction

```
ggplot(sinai_covid, aes(x = systolic_bp, y = diastolic_bp,
          color = facility)) +
  geom_point() +
  scale_color_brewer(palette = 3, direction = -1)
```

- Using manual scale

```
ggplot(sinai_covid,
       aes(x = systolic_bp, y = diastolic_bp,
           color = facility)) +
  geom_point() +
  scale_color_manual(values = c("red", "blue","darkgreen",
                                "orange", "pink"))
```
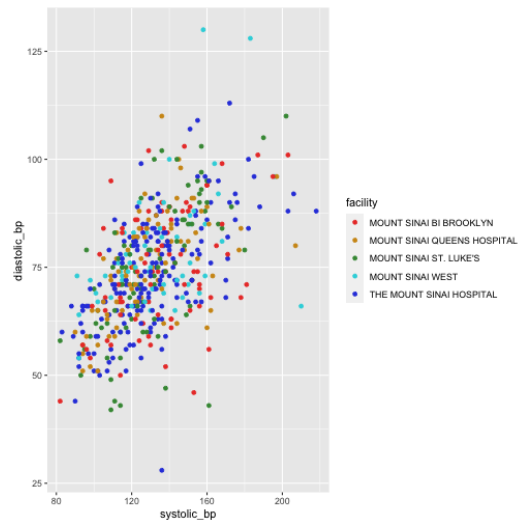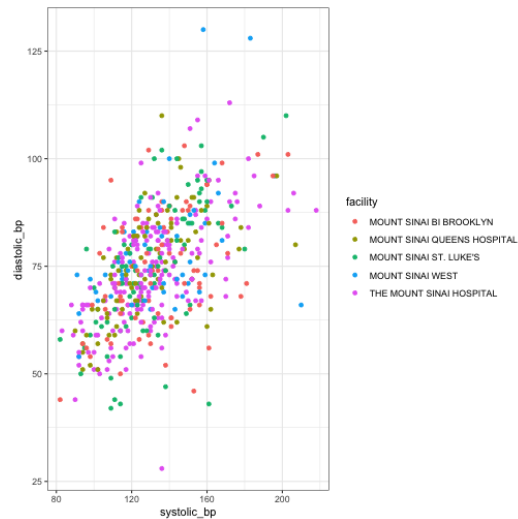
- Using manual scale

```
ggplot(sinai_covid,
       aes(x = systolic_bp, y = diastolic_bp,
           color = facility)) +
  geom_point() +
  scale_color_manual(values = c("#eb4034", "#d19617", "#429642",
                                "#31d5de","#3148de"))
```
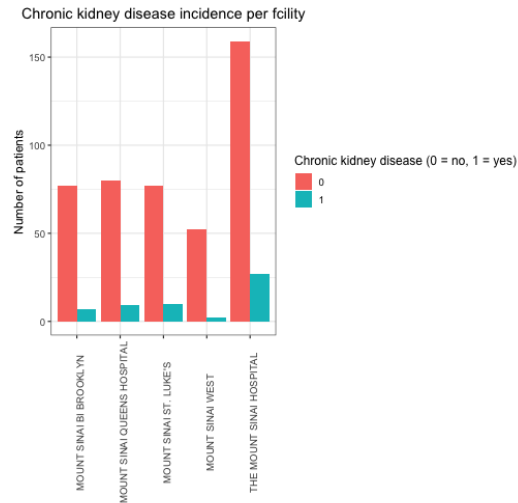

Look for "html color picker" on Google browser

themes

```
ggplot(sinai_covid,
       aes(x = systolic_bp, y = diastolic_bp,
           color = facility)) +
  geom_point() +
  theme_bw()
```

```
ggplot(kd,
       aes(x = facility, y = n, fill = factor(chronic_kidney_diseas
  geom_col(position = position_dodge()) +
  labs(title = "Chronic kidney disease incidence per fcility",
       x = "",
       y = "Number of patients",
       fill = "Chronic kidney disease (0 = no, 1 = yes)") +
  theme_bw(base_size = 12) +
  theme(plot.title = element_text(hjust = 0.5),
        axis.text.x = element_text(angle = 90))
```

Let's practice!

- Count the number of patients depending on their smoking status per etnic group
- Plot the data using vertical bars
- Add a plot title, axis titles and modify the legend title if needed
- Change the filling colors of the bars (use scale_fill_brewer or scale_fill_manual)
- Explore the available themes and use one of them
- Use the theme layer to modify the angle or size of the axis text if needed

Thanks!

Ilustration by Allison Horst